

Towards a ‘Machiavellian’ Theory of Emotional Appraisal¹

Paul E. Griffiths,
Department of History and Philosophy of Science,
University of Pittsburgh,
Pittsburgh,
PA 15232,
USA

pauleg@pitt.edu

1. Emotional Appraisal

The aim of appraisal theory in the psychology of emotion is to identify the features of the emotion-eliciting situation that lead to the production of one emotion rather than another². A model of emotional appraisal takes the form of a set of dimensions against which potentially emotion-eliciting situations are assessed. The dimensions of the emotion hyperspace might include, for example, whether the eliciting situation fulfills or frustrates the subject’s goals or whether an actor in the eliciting situation has violated a norm. Richard Lazarus’s well-known model of emotional appraisal has six dimensions, and the regions of the resulting hyperspace that correspond to particular emotions are summarized by Lazarus as the ‘core relational themes’ of those emotions. Anger, for examples, is elicited by the core relational theme ‘a demeaning offence against me and mine’, sadness by ‘having experienced an irrevocable loss’ and guilt by ‘having transgressed a moral imperative’ (Lazarus, 1991).

Dimensional appraisal models have traditionally been tested by asking people who have experienced a particular emotion to report on the appraisal process, or even by asking people to report on the relevance of certain dimensions of evaluation to certain emotions. It has therefore been alleged that appraisal theories are based, not on the reality of emotion processes, but rather on the image of those processes recorded in folk-wisdom. As the leading appraisal theorist Klaus Scherer puts it, some supposedly *psychological* studies may in fact ‘do little more than explicate the implicational semantic structures of our emotion vocabulary’ (Scherer, 1999, 655). This challenge to appraisal theory can be met in a number of ways, including prospective studies that manipulate situational factors relevant to the dimensions of appraisal and predict the resultant change in emotion, and studies that rely on objective measures of emotional behavior and physiology rather than on self-report.

¹ To appear in the proceedings of the Emotions, Evolution and Rationality conference, Kings College London, April 2002. In preparing this paper I am indebted to feedback from the audience at Kings and to the comments of Andrea Scarantino on an earlier draft.

² For a review of appraisal theories, see (Scherer, 1999).

The ongoing effort to test appraisal theories as theories of emotion, rather than as elucidations of folk wisdom, has led to a consensus that emotions do not walk in step with cognitive evaluation of the stimulus unless the notion of 'cognitive evaluation' is broadened to include sub-personal processes (Teasdale, 1999). Many appraisal theorists have come to accept that even such apparently conceptually complex appraisals as Lazarus's core relational themes can be assessed: 1. Without the information evaluated being available to other cognitive processes, 2. Before perceptual processing of the stimulus has been completed, and 3. Using only simple, sensory cues to define where the eliciting situation falls on the dimensions.

The existence of 'low-level' appraisal is demonstrated when emotional responses are independent of or conflict with explicit, reportable, action-guiding evaluations; when people are afraid of things that they know are not dangerous or disgusted by things they know to be healthy. Robert Zajonc's findings about 'affective primacy' are well known. He showed that subjects can form preferences for stimuli to which they have been exposed subliminally so that their ability to identify those stimuli remains at chance levels (Zajonc, 1980). Many results have since been obtained which confirm and extend this phenomenon. Arne Öhman and his collaborators have conditioned subjects to dislike angry faces and subsequently elicited the conditioned emotional response when those angry faces were masked by neutral faces so that subjects reported no conscious experience of the angry faces (Esteves & Öhman, 1993; Öhman, 1986). In another study, subjects were exposed to subliminal images of snakes, spiders, flowers and mushrooms. Although the subjects were unable to identify better than chance the stimuli they had been exposed to, subjects with pre-existing snake phobia showed elevated skin conductance responses to the snake images and subjects with spider phobia showed this response to the spider images (Öhman & Soares, 1994)³.

The original controversy aroused by Zajonc's concept of affective primacy concerned whether emotions involve a 'cognitive evaluation of the stimulus' (Lazarus, Coyne, & Folkman, 1984; Zajonc, 1984a, 1984b). It has become clear that this was not a helpful formulation, and that what is really at issue is whether the information processing that leads to an emotional response is separate from that which leads to paradigmatically cognitive processes such as conscious report and recall, and whether the two informational processes are different in kind. It is not necessary to deny, as Zajonc did, that low-level appraisal is a cognitive process. If emotional appraisal proceeds on several different *cognitive* levels, then traditional views about appraisal will still have to be very significantly modified. The traditional view, expressed here by Robert Solomon, is just too simple:

'all emotions presuppose or have as their preconditions, certain sorts of cognitions - an awareness of danger in fear, recognition of an offense in anger, appreciation of someone or something as lovable in love. Even the most hard-headed neurological or behavioral theory must take account of the fact that no matter what the neurology or the behavior, if a person is demonstrably ignorant of a

³ For a brief overview, see (Öhman, 2002).

certain state of affairs or facts, he or she cannot have certain emotions.’ (Solomon, 1993: 11).

The concept of ‘multi level appraisal’ combines the recognition 1) that emotions are elicited by information about certain aspects of the environment and that the emotional response is, in some sense, directed at those aspects of the environment and 2) that these states can occur at many cognitive levels, and an appraisal that leads to an emotion can be separate from, and can contradict, the appraisal of the same stimulus that is verbally reportable and integrated with the organisms other beliefs. The phobic subjects in the study just discussed were, in this sense, ‘demonstrably ignorant’ of the state of affairs that was the object of their emotional response. Often, of course, high-level appraisal and low-level appraisal march in step with one another, but this is not always the case, and the early advocates of affective primacy were fundamentally correct that even under normal conditions there are two (or more) processes going on. Paul Ekman embodies these ideas in his concept of an ‘automatic appraisal mechanism’ - a cognitive subsystem that is dedicated to determining whether a stimulus will elicit an basic emotion and which is able to operate independently of the cognitive systems that lead to conscious, verbally reportable appraisals of the same stimulus (Ekman, 1980; Griffiths, 1990).

The multi-level approach to appraisal has been solidly confirmed in the case of fear by the neuroscientist Joseph LeDoux (LeDoux, 1996; LeDoux, 1993). LeDoux distinguishes between ‘cognitive computations’ which yield information about stimuli and the relations between them, and ‘affective computations’ which yield information about the significance of stimuli for the organism and lead to physiological and behavioral responses appropriate to that significance. In fear, and possibly anger, key aspects of affective computation occur in the amygdala. The emotional evaluation of a stimulus can be driven by inputs at various levels of analysis. At a very early stage of perceptual processing, minimally processed data from thalamic sensory relay structures follows the ‘low road’ to the amygdala. This is the ultimate ‘quick and dirty’ route to rapid emotional response. Meanwhile, perceptual information follows a slower ‘high road’ to the visual, auditory, somatosensory, gustatory and olfactory cortices, projections from which to the amygdala allow responses to stimuli in a single, sensory modality. Lesions to these pathways inhibit emotional responses to stimulus features in the corresponding modalities. Finally, the amygdala receives inputs from brain regions associated with full-blown, polymodal, perceptual representations of the stimulus situation and with memory, allowing the emotional response to be triggered by complex, contextual features of the stimulus. However it is triggered, it is the final response in the amygdala that is associated with fear conditioning, and conditioned fear responses to simple sensory-perceptual stimuli have been shown to be resistant to extinction.

2. Appraisal and the philosophy of emotion

Appraisal theories are the closest scientific equivalents to the theories that have dominated philosophy of emotion since the 1960s. Philosophers have analyzed emotions in terms of their states of affairs appropriate to them (their ‘formal objects’) (Kenny, 1963), as evaluative judgments (Solomon, 1976), as evaluative judgments that cause

bodily arousal (Lyons, 1980), as feelings of comfort or discomfort directed towards an evaluative thought (Greenspan, 1988) and as the results of either true belief or uncertainty about emotion-inducing situations (Gordon, 1987). For these and many other authors, the central aim of a philosophical theory of emotion is to identify the content of an emotion - the actual or imagined state of affairs in the world to which the emotion has a semantic relationship. In her recent book *Upheavals of Thought: The intelligence of emotions*, Martha Nussbaum uses appraisal theory to capture the idea that emotions are 'intelligent responses to the perception of value' (Nussbaum, 2001, 1). She describes Richard Lazarus's appraisal theory as 'in all essentials the view of emotions I have defended in Chapter 1' (Nussbaum, 2001, 109). Nussbaum's treatment of the emotions in children and animals also makes use of something like multi-level appraisal. Animals as well as humans make the evaluative judgments that constitute emotions according to Nussbaum's theory, but they do so without self-conscious awareness and in such a way that the content of their judgments cannot be rendered in language without distortion. Nevertheless, she argues, emotion remains primarily an intentional phenomenon. Despite the existence of these low-level appraisal that cannot be expressed in language, '...emotions include in their content judgments that can be true or false, and good or bad guides to ethical choice' [Nussbaum, 2001 #6406, 1].

“What we need, in short, is a multifaceted notion of cognitive interpretation or seeing-as, accompanied by a flexible notion of intentionality that allows us to ascribe to a creature more or less precise, vaguer or more demarcated, ways of intending an object and marking it as salient.” (Nussbaum, 2001, 129)

Nussbaum's description of what is needed for an adequate account of emotional cognition is compelling, but levels of appraisal do not just differ *between* organisms. They also differ *within* a single organism. Multi-level models of emotional appraisal suggest that the same stimulus can be represented in several places in the human brain by different representations. Hence it is vital to understand not only what these multiple appraisals have in common, but also how they *differ* and how they *interact*. The existence of multiple representations in a 'hierarchical' emotional architecture (DeLancey, 2001) violates a key assumption of most philosophical reasoning about emotion, which is that emotional cognition manipulates emotional representations on the basis of their content, and thus that emotional processes can be explored via the semantic 'logic' of emotions. Multi-level models imply that how emotional and other representations interact, if they interact at all, depends on details of cognitive architecture as well as on the content of the representations. This architecture, of course, cannot be determined by studying the conceptual relations between the contents of emotional representations. Phobias and affective primacy phenomena reveal the architecture of the emotion system by showing what happens when one level of appraisal operates without the others. They reveal that certain information, such as partially analyzed visual data, is available to one level of appraisal and not to another. Another insight is provided by people with 'flattened affect', who are apparently able to carry out high-level appraisal but not low-level appraisal and so do not experience the physiological components of normal emotional response. The possibility of flattened affect without intellectual impairment reveals that only low-level

appraisal has direct connections to the effector systems for the automated components of rapid emotional response.

Normal human emotion involves several subsystems that interact, and interact with other cognitive subsystems, in ways that reflect the particular cognitive architecture in which they are embedded. This renders the idea that emotions are intentional states whose identity is determined by their representational content seriously incomplete, since the same state of affairs is multiply represented. If low-level emotional appraisals are not merely separate from high-level appraisals but different in kind, that idea becomes still more inadequate. Nussbaum (p.127-8) notes that our language may not adequately express the content of appraisal processes in animals and perhaps also those of infants. I would add adult low-level appraisal processes to this list. Many of the fine-grained semantic distinctions we make in ascribing content to thoughts may fail to get a grip on representational states with more coarse-grained semantics. Neither 'aerial predator' nor 'dark thing overhead of such-and-such apparent size and moving thusly' should really be taken as a literal rendition of the 'what the crow thinks' before exhibiting aerial predator-avoidance behavior, although both point to something meaningful (see below). Nussbaum suggests that we can cope with this phenomenon using a 'flexible notion of intentionality'. Since this is supposed to allow us to identify what is in common between animal and human emotion her idea is presumably that there is some degree of isomorphism between the way in which high-level representations relate to one another on the basis of their content and the way in which low-level representations relate to one another on the basis of their 'content'. The next two points suggest that this hope will not be borne out.

There are likely to be radical differences between the representational states involved in low-level and high-level appraisal. Ruth Millikan has suggested that mental representations in simple organisms may unite the functions of beliefs and desires in a single, undifferentiated functional role. Low-level appraisal in humans seems to manifest the same 'collapse of the attitudes'. Consider the low-level appraisal of the core relational theme 'a demeaning offence to me and mine' that presumably occurs when a soccer player is dribbling the ball down the field, another player grabs his jersey causing him to lose the ball, and the first player turns *angrily* towards the second. I suggest that it is misleading to say that the relevant brain region *believes* that the core relational theme has been instantiated. Beliefs are mental states that represent how things are and which produce action in conjunction with desires - representations of how the world should be. But in low-level appraisal for anger there is no question as to what action will be taken. The frustrated player in our example will orient to the stimulus, produce the pan-cultural facial expression of anger and undergo physiological changes to prepare them for aggressive action. The 'affective computation' in this example is simultaneously the belief that the world is a particular way and the intention to act in a particular way. Likewise for the better understood case of affective computing of fear in the amygdala, and, presumably, for any emotion that has a clear behavioral signature and can be induced to exhibit affective primacy. I suggest that it is simply misleading to describe low-level appraisal as evaluative judgment, or using any other locution derived from a psychology that presumes a fundamental distinction between data and goals. Instead,

low-level emotional appraisal seems to involve action-oriented representation. I discuss below the possibility of capturing the content of such representations using the notion of 'affordances'.

Another way in which low-level emotional appraisal may differ from high-level is in terms of the narrow inferential role imposed on low-level representations by the task-specific architecture in which they occur. The inferential role of these representations is impoverished in three ways. First, low-level appraisal processes do not have access to most of what is represented elsewhere in the brain, which is why knowledge that the cockroach in my drink has been completely sterilized does not eliminate the disgust response. Hence many inferences that would seem to follow from the content we ascribe to this low-level appraisal - in Lazarus's theory, something like 'I am taking in or being too close to an indigestible object' - do not actually follow for subjects because they cannot recombine that content appropriately with their other contentful states. Secondly, the actual processes of affective computing (as opposed to their final output) are not available for inspection by other cognitive sub-systems. Once again, architectural barriers to information flow block inferences that follow from what otherwise seems the natural content to ascribe to those states. Finally, the inferential principles used in affective computing are not truth-preserving, but heuristically survival-enhancing. It simply does not follow from the fact that I have been poked hard and unexpectedly in the small of my back that I have suffered 'a demeaning offence to me and mine' but the automatic appraisal mechanism for anger will reliably draw that conclusion⁴.

There is no scientific puzzle about the nature of information processing in affective computing. The forms of inferential impoverishment I have described all make good psychological and evolutionary sense. But there is a considerable philosophical puzzle about how to ascribe conceptual content to representational states in an isolated cognitive subsystem of this kind. If the concepts that figure in the content ascribed to a representation do not have their usual inferential role, then what is meant by attributing that content? The very idea that the state has conceptual content is thereby called into question⁵. The actual role of the representations involved in low-level appraisal and the inferential role of the content-sentences with which we describe those appraisals strongly suggests that, in this role at least, appraisal theories simply are not theories of cognitive content.

⁴ Some Evolutionary Psychologists would say that the appraisal mechanism has innate knowledge that this cue reliably predicts conspecific aggression, or did so in some ancestral environment (Tooby & Cosmides, 1992). I take this to be only rhetorically different from the claim that the appraisal mechanism consistently makes certain logically invalid inferences.

⁵ Gareth Evans famously argued that for a state to have conceptual content its elements must be able to be recombined appropriately with all the other conceptual elements in representational states of the same organism (the 'generality constraint') (Evans, 1982). See also Stephen Stich's discussion of 'sub-doxastic' mental representations. The 'beliefs' of low-level appraisal mechanisms resemble those of Stich's 'Mrs T', who remembered that President McKinley was assassinated but was unsure if he was alive or dead (Stich, 1983).

3. What is appraisal theory a theory of?

The idea that a low-level appraisal and a high-level appraisal can be *the same appraisal* poses the question of what these two processes have in common? According to multi-level appraisal theories, what they have in common is expressed by the single appraisal that can be made at the various levels, an appraisal such as ‘There has been a demeaning offence to me and mine’. I have argued that this shared appraisal cannot be understood as the shared conceptual content of the various appraisal processes, because it is problematic say in what sense low-level appraisal has conceptual content at all. Given the collapse of the attitudes and the impoverishment of inferential roles of representational states in low-level appraisal it makes sense to describe those appraisal processes as ‘sub-conceptual’. Instead, I suggest that multi-level appraisal theories are essentially *ecological* theories - theories of the significance of the environment for the organism. An appraisal hyperspace identifies the aspects of the environment that the organism tracks in order to produce adaptive behavior in that environment. The relationship between the concepts that figure in the appraisal theory’s representation of the environment and the way the organism actually tracks those aspects of the environment may be very indirect. Lazarus describes the appraisal associated with shame as ‘I have failed to live up to an ego-ideal’ but low-level appraisal leading to shame may be driven by simple cues that indicate my social standing relative to the person with whom I am interacting⁶. The fact that different appraisal levels are using different surrogates to track the same ecologically significant aspects of the environment helps explain why the different levels sometimes fail to coincide on a common appraisal.

There is an important sense in which a particular account of the content of emotional appraisal can live a double life, first as a general ecological theory applicable to all appraisal levels and second as a theory of the actual conceptual content of the representational states being processed in high-level appraisal. An appraisal theory can fulfill the first role very well whilst failing in the second role. In fact, abstracting away from the details of any particular psychological process in order to better fulfill the first role is likely to produce a worse theory of the actual psychological process of high-level appraisal. But if an appraisal theory is not primarily treated as a theory of ‘what the person is thinking’ then it becomes possible to recognize that even high-level appraisal may not present the stimulus to the subject under the concepts that figure in the best theoretical representation of the appraisal hyperspace. For example, if high-level appraisal for shame involves moral concepts from the local cultural milieu, such as ‘dishonour’, whilst low-level appraisal uses simple cues to assess relative social status (see fn.5 above), then the core relational theme ‘I have failed to live up to an ego-ideal’ is not an accurate rendition of either the conceptual content of the high-level appraisal or the ‘content’ of the low-level appraisal. It may, however, capture a common aspect of eliciting situations that can be tracked at many levels. In earlier work on the philosophical theories of the content of emotions I discussed the problem of relating the very abstract

⁶ Anthropologist Daniel Fessler has argued that shame is built out of an ancient behavioral system that leads to aversive feelings in social interactions where one’s low relative social status is made salient and that this basic system continues to operate alongside forms of shame that tie the emotion to concepts of responsibility and norm-violation (Fessler, 1999).

‘analyses’ of emotions like fear, anger and envy found in the philosophical literature to the concrete thoughts about particular objects that are present in an individual’s consciousness on a particular occasion of emotion (Griffiths, 1989; 1997, 38-43). The present suggestion is that greater psychological realism about the conceptual content of high-level appraisal processes can be achieved by recognizing that the general gloss put on such thoughts in an appraisal theory is an ecological description of the significance of the aspects of the environment that the organism is tracking and not necessarily a description of how the organism represents those aspects of the environment to itself.

The idea that appraisal theories are ecological theories of the significance of the environment to the organism immediately suggests a teleosemantic treatment of the content of emotional appraisals, a treatment that it might be hoped would explain how appraisals at all levels could have the same conceptual content⁷. The idea would be that appraisals at all levels for, e.g. fear, have the same content - namely, ‘I am facing an immediate, concrete, and overwhelming physical danger’ - because they have all evolved to be triggered by that circumstance. But this proposal does not take the idea of assigning conceptual content to a mental representation seriously enough. Teleosemantic programs for naturalistic semantics see the fact that some biological systems have the evolutionary function of responding to a state of the environment as the fundamental way in which ‘aboutness’ gets to be a part of the natural world. The sort of genuine intentionality that characterizes human thought and language is something that can be explained as a sophisticated evolutionary development of this fundamental phenomenon. But that does not mean that any biological system with a broadly representational function is comparable to those distinctive biological systems like the human mind that achieve full-blown intentional representation. As I suggested above, low-level appraisal resembles just those sorts of proto-intentional representational processes that have been the main focus of the teleosemantics literature to date - processes like that by which a frog detects a passing fly and strikes at it (Lettvin, Maturana, & McCulloch, 1959). The idea that all levels of appraisal have the same ‘teleocontent’ comes down to nothing more than what I have already argued, namely, that a multi-level appraisal theory describes the ecological significance of the aspects of the environment that the organism is tracking by varied cognitive means.

Treating appraisal theory as an ecological theory, as opposed to a theory of teleocontent, has the advantage that it does not lump together the biological functions of the various appraisal levels for an emotion as ‘representing the core relational theme’, as if they were redundant ways to perform the same task. It should be clear that low-level appraisal systems are not merely a hangover from the evolutionary past, performing the same functions in primitive ways. Rather, the ways in which low-level appraisal conflict with high-level appraisal is part of the normal functioning of these systems. Low-level appraisal ensures that fundamental adaptive actions and action preparedness is engaged early and with a low rate of false-negative appraisals. High-level appraisal makes fuller use of human cognitive abilities to make an accurate appraisal that will engage flexible,

⁷ For teleosemantic accounts of mental content, see (Dretske, 1988; Millikan, 1984; Papineau, 1987). Jesse Prinz has a specific discussion of teleosemantics and the emotions, but his immediate concern is with what affective feelings represent to the organism (Prinz, Forthcoming).

longer-term coping strategies. So, rather than being redundant mechanisms performing one function, low-level appraisal and high-level appraisal performs separate and complimentary functions in relation to the same ecologically significant aspects of its environment - dangers, conspecific challenges, and so forth. It is useful to see what low-level and high-level appraisal have in common, but they are not the same thing.

4. Appraisal and Affordance

I have argued that low-level appraisal involves 'action oriented' representations in which the functional roles of belief and desire are not distinct. This makes it natural to describe the ecologically significant features of the environment that the appraisal process is tracking as 'action affordances' - the fact that the environment offers a certain opportunity for action. This approach may perhaps be useful for high-level appraisal as well as low-level appraisal. Nico Frijda has argued that the result of emotional appraisal is a set of 'action tendencies' - the emotion presents the environment to the subject as affording certain possibilities for action (Frijda, 1986).

The concept of an affordance derives from J.J Gibson's theory of, 'ecological perception'. Philosophical discussion of Gibson has focused on his 'direct realism' and rejection of perceptual representations. The concept of an affordance, however, is logically independent of direct realism. In a recent paper Andrea Scarantino has presented an analysis of the affordance concept and pointed out the theoretical value of considering affordances as objects of perception, independent of any particular theory of the perceptual process (Scarantino, In Press). Affordances are organism-relative dispositional properties of the environment. The surface of water, for example, affords mosquito larvae the possibility of hanging from it so as to keep their breathing tube exposed to the air. If the mosquito larva acts appropriately, then this disposition will be realized in an act of hanging from the surface. The disposition arises from a conjunction of properties of the larvae and their environment. Water does not afford this possibility to human infants. Adding petrol to the water reduces the surface tension and destroys the disposition.

Scarantino makes two, important, orthogonal distinctions between types of affordances. In some cases the eliciting condition guarantees the manifestation of the disposition, while in others it only makes it probable. First, therefore, he distinguishes deterministic ('surefire') affordances from probabilistic affordances. Secondly, Scarantino distinguishes 'goal-affordances' and 'happening affordances'. Goal-affordances are manifested by an organism successfully performing a goal-directed behavior - the environment affords the performance of that behavior. A member of the opposite sex, for example, offers a probabilistic goal-affordance of mating. Happening-affordances are manifested by the organism undergoing a passive change. A branch of dubious strength presents an orang-utang with a probabilistic happening-affordance of falling. These distinctions can both be applied to the properties detected in emotional appraisal processes.

Standard presentations of appraisal make emotions appear primarily perceptual. An emotion is the recognition that the world is a certain way. Representing the content of

appraisals as affordances brings out the action-oriented nature of emotion. Shame for example, has the core relational theme that 'I have failed to live up to an ego ideal'. But shame also has a characteristic set of behaviors that serve a vital social-communicative function (Darwin, 1872; Fessler, 1999). Arguably, the fact that the environment is shaming to the subject is a goal-affordance - the environment affords acceptance of the other's dominance as a social-interaction strategy. In some cases, the action-oriented nature of emotion is already present in the 'coping' dimensions of appraisal models and is simply not highlighted in the usual summaries of those models. The core relational theme of anger, for instance, is 'a demeaning offence against me and mine', but what this actually means in Lazarus's model includes the organism's recognition that aggression is a viable coping strategy and that aggression may help achieve the organism's goals (Lazarus, 1991). In anger, therefore, what we detect is a goal-affordance for angry behavior. Conceiving of anger in these terms makes sense of an intriguing result from a retrospective self-report study of the causes of emotion by Nancy Stein and colleagues. They found that an important factor in predicting the occurrence of anger rather than sadness as the response to a loss was possibility of obtaining restitution or compensation for the loss (Stein, Trabasso, & Liwag, 1993). Sadness, from this perspective, is an unusual emotion, since, if the core relational theme really is 'I have experienced an irrevocable loss' then sadness is a deterministic happening-affordance, or something about which nothing can be done! This poses an obvious question about the adaptive utility of the emotion.

5. Machiavellian appraisal?

In a recent paper I raised the possibility that some aspects of the appraisal process might be 'strategic' or 'Machiavellian' (Griffiths, In Press). Organisms may be evaluating not just the significance of what has happened, but also the likely outcome of responding to that occurrence with one emotion or another. The concept of 'Machiavellian intelligence' has moved to center stage in recent discussions of the evolution of human cognition (Byrne & Whiten, 1988; Whiten & Byrne, 1997). Human intelligence is 'Machiavellian' to the extent that the evolutionary forces which shaped it arose from social competition within primate groups. Machiavellian intelligence is the result of an intra-species arms race in which increased intelligence at the level of the population merely raises the bar for success at the individual level. Like more traditional conceptions of intelligence, emotional intelligence is Machiavellian in a general sense simply to the extent that it has evolved to mediate social competition.

It seems to be common ground that the *expression* of emotion is strategic, and so, if it has an evolutionary history, is Machiavellian in the above sense. The role of social context in determining emotional expression has long been recognized through Paul Ekman's concept of a 'display rule' (Ekman, 1971, 1972) and has been more heavily emphasized by recent 'transactionalist' accounts of emotion (Fridlund, 1994; Russell & Fernández-Dols, 1997). The contextual factors that predict whether an emotion is expressed are of the sort that are likely to have been significant in human evolution - factors such as conformity to group standards and the status of the individual in the group - and so the

sensitivity of emotional expression to such factors is very plausibly part of our evolved social competence. This need not imply, of course, that the specific rules to which individuals conform in one culture or another can be explained in evolutionary terms. Evolution can be equally relevant when the task is to understand how cultures generate their patterns of difference from a shared developmental system. The generic fact that there are display rules, for example, is very likely to have an evolutionary explanation even if these rules differ radically across cultures.

If the Machiavellian expression of emotions is common ground amongst evolutionary theorists of emotion, the idea that the actual production of emotion is influenced by 'strategic' considerations remains more controversial (Ekman, 1997). I have embodied this idea in the '*Machiavellian Emotion Hypothesis*', which suggests that emotional appraisal is sensitive to cues that predict the value to the emotional agent of responding to the situation with a particular emotion, as well as cues that indicate the significance of the stimulus situation to the agent independently of the agent's response (Griffiths, In Press). Put in the language of appraisal theories, the hypothesis is that the appraisal hyperspace has 'strategic' dimensions. The organism evaluates the stimulus situation not only along multiple dimensions that assess the significance of what has happened, but along dimensions that assess the significance of what will likely happen if the emotion is produced. In current appraisal models, some strategic considerations are embodied in 'coping' related dimensions. Putting the hypothesis in terms of affordances, it states that emotional appraisal detects goal-affordances.

There is some suggestive evidence of Machiavellian appraisal in primates. The neural connectivity of the amygdala in primates is consistent with the hypothesis that its activity is facilitated or inhibited by the perception of social contextual cues (Emery & Amaral, 2000: 167). Behavioral studies of primates also support the idea that adults assess the likely outcome of the emotional interaction when producing emotional behavior. William Mason reports that rhesus macaques deprived of social contact as infants produce grossly normal facial behaviors expressive of fear (grimace), friendliness (lipsmacking) and threat (threat face) (Mason, 1985). What seems to be lacking in these animals is an ability to utilize these facial expressions to manage their relationships with other monkeys. They are, as it were, emotionally clumsy. Mason explains these results in terms of the role of social experience in elaborating complex eliciting conditions for emotional behavior. Infant monkeys begin by producing these behaviors in response to relatively simple, context independent stimuli. Later on, "As a result of functional elaborations, refinements, and transformations of the schemata [*of elicitors for expressive behavior*] present in early infancy, experience creates new sources of social order, new possibilities for the regulation and control of social life." (Mason 1985: 147). The monkeys, in other words, learn to produce the same emotional behaviors in response to subtler aspects of social context, and by doing so are able to manage their social interactions with other animals. It is possible to argue that these findings are accounted for purely by evolved mechanisms for the Machiavellian *expression* of emotions - subordinate Rhesus monkeys feel angry when dominant monkeys search their cheek pouches for food, but never display this anger. I would argue, however, that this hypothesis adds an unnecessary layer of complexity to account for the behavior.

What would a Machiavellian theory of human emotions look like? Numerous emotion theorists have suggested that emotions may be self-serving, occurring not when the situation objectively warrants the judgment embodied in the emotional appraisal, but rather when it suits the agent to interpret the situation in this light. Jean Paul Sartre famously takes this view (Sartre, 1962). According to Sartre, for example, anger ascribes to a person the property of being hateful precisely because he stands between the agent and the satisfaction of her desires. That is the difference between the emotion of anger and rational coping with the conflicting needs of others. Emotions as Sartre describes them are intrinsically pathological - a form of bad faith in which people reject reality out of mental weakness. But the central insight of his theory is independent of this value-judgment. It is simply that people can use emotions to view the world in a light that is psychologically more rewarding to us than other possible interpretations. Highly adaptive versions of this process are described in the literature on 'emotional intelligence', such as using an emotional reinterpretation of the situation to motivate oneself (Salovey, Bedell, Detweiler, & Mayer, 2000). From the perspective of the emotional intelligence literature, Sartre's work seems like an insightful account of human psychology marred by the French philosophers penchant for calling a spade a conspiracy against the soil.

6. Conclusion

Emotional appraisal happens at more than one level. Low-level appraisals involve representations that are semantically coarse-grained, fuse the functional roles of belief and desire and have impoverished inferential roles, making it best to think of them as sub-conceptual. Multi-level theories of emotional appraisal are thus best conceived, not as theories of the actual conceptual content of emotional appraisals, but as ecological theories that identify the aspects of the environment that appraisal processes are tracking using diverse cognitive means. These aspects of the environment are what the environment 'affords' the organism. Some of these affordances are 'goal-affordances' - possibilities for future action. This perspective on emotional appraisal lends support to the idea that emotional appraisal is in part 'Machiavellian' or 'strategic'. Organisms take into account the payoffs resulting from an emotional response when determining whether the eliciting situation 'warrants' that emotion.

References

- Byrne, R. W., & Whiten, A. E. (1988). *Machiavellian Intelligence*. Oxford, New York: Oxford University Press.
- Darwin, C. (1872). *The Expressions of Emotions in Man & Animals* (1st ed.). New York: Philosophical Library.
- DeLancey, C. (2001). *Passionate Engines: What emotions reveal about mind and artificial intelligence*. New York, Oxford: Oxford University Press.
- Dretske, F. (1988). *Explaining Behaviour*. Cambridge, MA: Bradford/MIT.

- Ekman, P. (1971). Universals and Cultural Differences in Facial Expressions of Emotion. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation 4*. (pp. 207-283). Lincoln, Nebraska: University of Nebraska Press.
- Ekman, P. (1972). *Emotions in the Human Face*. New York: Pergamon Press.
- Ekman, P. (1980). Biological & cultural contributions to body & facial movement in the expression of emotions. In A. O. Rorty (Ed.), *Explaining Emotions* (pp. xxx-xxx). Berkeley: University of California Press.
- Ekman, P. (1997). Expression or communication about emotion. In N. Segal & G. L. Weisfeld & C. C. Weisfeld (Eds.), *Genetic, Ethological and Evolutionary Perspectives on Human Development: Essays in Honor of Dr Daniel G. Freedman* (pp. xxx-xxx). Washington, D.C: American Psychiatric Association.
- Emery, N. J., & Amaral, D. G. (2000). The role of the amygdala in primate social cognition. In R. Lane & L. Nadel (Eds.), *The Cognitive Neuroscience of Emotion* (pp. 156-191). New York: Oxford University Press.
- Esteves, F., & Öhman, A. (1993). Masking the face: recognition of emotional facial expression as a function of the parameters of backward masking. *Scandinavian Journal of Psychology*, 34(1-18).
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
- Fessler, D. M. (1999). Toward an understanding of the universality of second-order emotions. In A. L. Hinton (Ed.), *Biocultural approaches to emotions* (pp. 74-116). Cambridge: Cambridge University Press.
- Fridlund, A. (1994). *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press.
- Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.
- Gordon, R. M. (1987). *The Structure of Emotions*. Cambridge: Cambridge University Press.
- Greenspan, P. (1988). *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge.
- Griffiths, P. E. (1989). The Degeneration of the Cognitive Theory of Emotion. *Philosophical Psychology*, 2 (3), 297-313.
- Griffiths, P. E. (1990). Modularity & the Psychoevolutionary Theory of Emotion. . *Biology & Philosophy*, 5, 175-196.
- Griffiths, P. E. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Griffiths, P. E. (In Press). Basic Emotions, Complex Emotions, Machiavellian Emotions. In A. Hatzimoysis (Ed.), *Proceedings of the Royal Institute of Philosophy, 2001* (pp. xxx-xxx).
- Kenny, A. (1963). *Action, Emotion & Will*. London: Routledge & Kegan Paul.
- Lazarus, R. S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.
- Lazarus, R. S., Coyne, J. C., & Folkman, S. (1984). Cognition, emotion & motivation: doctoring Humpty Dumpty. In K. Scherer & P. Ekman (Eds.), *Approaches to Emotions* (pp. 221-237). Hillsdale, New Jersey: Erlbaum.
- LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.
- LeDoux, J. E. (1993). Emotional networks in the brain. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 109-118). New York: Guildford Press.

- Lettvin, J. Y., Maturana, H. R., & McCulloch, W. S. P., W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of Institute of Radio Engineers*, 11, 230-255.
- Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.
- Mason, W. A. (1985). Experiential influences on the development of expressive behaviors in Rhesus monkeys. In G. Zivin (Ed.), *The Development of Expressive Behavior* (pp. 117-152). New York: Academic Press.
- Millikan, R. G. (1984). *Language, Thought & Other Biological Categories*. Cambridge, Mass.: M.I.T. Press.
- Nussbaum, M. C. (2001). *Upheavals of thought : the intelligence of emotions*. Cambridge, New York: Cambridge University Press.
- Öhman, A. (1986). Face the beast and fear the face: animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*, 23, 123-145.
- Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science*, 11(2), 62-66.
- Öhman, A., & Soares, J. J. F. (1994). Unconscious anxiety: phobic responses to masked stimuli. *Journal of Abnormal Psychology*, 102, 121-132.
- Papineau, D. (1987). *Reality and Representation*. NY: Blackwells.
- Prinz, J. (Forthcoming). *Emotional Perception*. Oxford: Oxford University Press.
- Russell, J. A., & Fernández-Dols, J. M. (1997). *The Psychology of Facial Expression*. Cambridge: Cambridge University Press.
- Salovey, P., Bedell, B. T., Detweiler, J. B., & Mayer, J. D. (2000). Current Directions in Emotional Intelligence Research. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (2 ed., pp. 5-4-520). New York: Guildford Press.
- Sartre, J. P. (1962). *Sketch for a Theory of the Emotions* (P. Mairet, Trans.). London: Methuen.
- Scarantino, A. (In Press). Affordances Explained. *Philosophy of Science (Supplement)*.
- Scherer, K. R. (1999). Appraisal Theory. In T. Dalgleish & M. J. Power (Eds.), *Handbook of Emotion and Cognition* (pp. 637-663). Chichester: New York.
- Solomon, R. (1976). *The Passions*. New York: Doubleday.
- Solomon, R. C. (1993). The philosophy of emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 3-15). New York: Guildford.
- Stein, N. L., Trabasso, T., & Liwag, M. (1993). The representation and organization of emotional experience: unfolding the emotion episode. In M. Lewis & J. M. Haviland (Eds.), *Handbook of Emotions* (pp. 279-300). New York: Guildford Press.
- Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, U.K.: M.I.T. Press.
- Teasdale, J. D. (1999). Multi-level Theories of Cognition-Emotion Relations. In T. Dalgleish & M. J. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 665-681). Chichester: John Wiley and sons.
- Tooby, J., & Cosmides, L. (1992). The Psychological Foundations of Culture. In J. H. Barkow & L. Cosmides & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 19-136). Oxford and New York: Oxford University Press.

- Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge: Cambridge University Press.
- Zajonc, R. B. (1980). Feeling & thinking: preferences need no inference. *American Psychologist*, 35, 151-175.
- Zajonc, R. B. (1984a). On the primacy of affect. In K. Scherer & P. Ekman (Eds.), *Approaches to Emotion* (pp. 259-270). Hillsdale, N.J: Lawrence Erlbaum Assoc.
- Zajonc, R. B. (1984b). The interaction of affect and cognition. In K. Scherer & P. Ekman (Eds.), *Approaches to Emotions* (pp. 239-246). Hillsdale, NJ: Lawrence Erlbaum Associates.